

# REGULAR TRANSITION MATRICES

JOSEPH TOBIN

## 1. INTRODUCTION

In this paper, we explore properties of regular transition matrices using the guidance of Friedberg, Insel, and Spence [FIS03]. We will start off with basic definitions and then dive into theories, all of which are shown with sufficient proofs. Then, we will deduce the relevant theorems and develop a whole swath of properties for square regular transition matrices. Finally, we will explore applications of these theorems.

We rely heavily on results shown in previous parts of section 5.3 in addition to results shown section in 5.1 and 5.2 of [FIS03]. Additionally, previous knowledge of Jordan Canonical Form is assumed.

## 2. BASIC DEFINITIONS AND NOTES

**Definition 2.1** (Matrix Limit). Let  $L, A_1, A_2, \dots$ , be  $n \times p$  matrices having complex entries. The sequence  $A_1, A_2, \dots$  is said to **converge** to the  $n \times p$  matrix  $L$ , called the **limit** of the sequence, if  $\forall 1 \leq i \leq n$  and  $\forall 1 \leq j \leq p$

$$\lim_{m \rightarrow \infty} (A_m)_{ij} = L_{ij}.$$

In this case, we write

$$\lim_{m \rightarrow \infty} (A_m) = L.$$

**Definition 2.2** (Transition Matrix, Probability Vector). We call an  $n \times n$   $M$  a **transition matrix** if it contains only nonnegative entries and all of its columns sum to 1. We call a column vector  $P$  a **probability vector** if it contains only nonnegative entries that sum to 1.

*Remark 2.3* (Theorem 5.15). (1)  $M$  is a transition matrix if and only if  $M^t u = u$ .

(2)  $v$  is a probability vector if  $u^t v = (1)$ .

(3) The product of two transition matrices is a transition matrix.

(4) The product of a transition matrix and a probability vector is a probability vector.

**Definition 2.4** (Regular). A transition matrix  $M$  is called regular if there exists an  $s \in \mathbb{N}_{>0}$  such that  $M^s$  has only positive entries.

**Definition 2.5** (Row Sum, Column Sum). Let  $A \in M_{n \times n}(C)$ .

$\forall 1 \leq i, j \leq n$ , define  $\rho_i(A)$  to be the sum of the absolute values of the entries of row  $i$  of  $A$  or

$$\rho_i(A) = \sum_{j=1}^n |A_{ij}|$$

and define  $\nu_j(A)$  to be the sum of the absolute values of the entries of column  $j$  of  $A$  or

$$\nu_j(A) = \sum_{i=1}^n |A_{ij}|$$

Then the **row sum** of  $A$  or  $\rho(A)$  and the **column sum** or  $\nu(A)$  are defined as:

$$\begin{aligned} \rho(A) &= \max\{\rho_i(A) \mid \forall 1 \leq i \leq n\} \\ \nu(A) &= \max\{\nu_j(A) \mid \forall 1 \leq j \leq n\} \end{aligned}$$

### 3. THEOREMS

**Lemma 1** (Section 6 Exercise 15(b)). Suppose  $V$  is an inner product space. Show  $\|x + y\| = \|x\| + \|y\|$  if and only if  $x$  is a non-negative multiple of  $y$ .

*Proof.*  $\rightarrow$  Suppose  $x = cy$  Then

$$\|x + y\| = \|cx + y\| = \|(c + 1)y\| = |c + 1|\|y\| = (c + 1)\|y\| = |c|\|y\| + \|y\| = \|cy\| + \|y\| = \|x\| + \|y\|$$

$\leftarrow$  We know by the proof of the triangle inequality (Theorem 6.2 (d) [FIS03]) that  $\|x + y\|^2 = (\|x\| + \|y\|)^2$  only when

$$\operatorname{Re}(\langle x, y \rangle) = |\langle x, y \rangle| = \|x\| * \|y\|$$

where  $\operatorname{Re}(x)$  is the real part of a complex number  $x$ .

But this only holds if  $x = cy$  for some nonnegative  $c$ .

□

*Remark 3.1.* Then  $\forall n \in \mathbb{N}$ , consider  $\|\sum_{i=1}^n x_i\|$ . Then by the previous lemma,

$$\|\sum_{i=1}^n x_i\| = \|\sum_{i=1}^{n-1} x_i + x_n\|$$

. Then we can repeat to get

$$\|\sum_{i=1}^n x_i\| = \|\sum_{i=1}^{n-1} x_i\| + \|x_n\| = \|\sum_{i=1}^{n-2} x_i\| + \|x_{n-1}\| + \|x_n\|$$

And so on until

$$\|\sum_{i=1}^n x_i\| = \sum_{i=1}^n \|x_i\|$$

**Theorem 2** (Theorem 5.18 pg. 298). Let  $A \in M_{n \times n}(C)$  be a matrix in which each entry is positive and let  $\lambda$  be an eigenvalue of  $A$  such that  $|\lambda| = \rho(A)$ . Then  $\lambda = \rho(A)$  and  $\{u\}$  is a basis for  $E_\lambda$ , where  $u \in C^n$  is the column vector in which each coordinate equals 1.

*Proof.* First, note that because  $A$  has all positive values and  $u$  has all positive values, then  $Au$  has all positive values. Therefore, because  $\lambda u = Au$ ,  $\lambda u$  has all positive values. Since  $u$  has each coordinate as 1, we can conclude  $\lambda > 0$  and thus  $\lambda = |\lambda| = \rho(A)$

Now we want to show that  $\{u\}$  is a basis for  $E_\lambda$ . To show this, we are going to use the fact that  $|\lambda| = \rho(A)$  to derive several equalities giving us information about  $A$ .

First, let  $v$  be an eigenvector of  $A$  corresponding to  $\lambda$  with coordinates  $v_1, v_2, \dots, v_n$ . Then let  $v_k = \max(|v_1|, |v_2|, \dots, |v_n|)$  and  $b = |v_k|$ .

Then

$$|\lambda| b = |\lambda| |v_k| = |\lambda v_k|$$

But if  $\lambda$  is an eigenvalue of  $A$ , then  $Av = \lambda v$  and thus  $\forall 1 \leq i \leq n$ ,  $\lambda v_i = \sum_{j=1}^n A_{ij} v_j$ . Thus

$$|\lambda v_k| = \left| \sum_{j=1}^n A_{kj} v_j \right|$$

By the triangle inequality and then absolute value multiplication rules,

$$\left| \sum_{j=1}^n A_{kj} v_j \right| \leq \sum_{j=1}^n |A_{kj} v_j| = \sum_{j=1}^n |A_{kj}| |v_j|$$

Since we know  $b = |v_k| \geq |v_j| \forall 1 \leq j \leq n$  and similarly  $\rho(A) \geq \rho_i(A) \forall 1 \leq i \leq n$ , we know

$$\sum_{j=1}^n |A_{kj}| |v_j| \leq \sum_{j=1}^n |A_{kj}| b = b \sum_{j=1}^n |A_{kj}| = b \rho_k(A) = \rho_k(A) b \leq \rho(A) b$$

But since we know  $|\lambda| = \rho(A)$ , we know the three inequalities used above are actually equalities.

(1)

$$|\sum_{j=1}^n A_{kj} v_j| = \sum_{j=1}^n |A_{kj} v_j|$$

(2)

$$\sum_{j=1}^n |A_{kj}| |v_j| = \sum_{j=1}^n |A_{kj}| b$$

(3)

$$\rho_k(A)b = \rho(A)b$$

But now we can use this to show that  $\{u\}$  is a basis for  $E_\lambda$

□

**Corollary 2.1** (Corollary 1 pg. 299 [FIS03]). Let  $A \in M_{n \times n}(C)$  be a matrix in which each entry is positive and let  $\lambda$  be an eigenvalue of  $A$  such that  $|\lambda| = \nu(A)$ . Then  $\lambda = \nu(A)$  and  $E_\lambda$  has dimension 1.

*Proof.* Consider  $A^t$  and let  $E_\lambda, E_{\lambda'}$  be the eigenspaces of  $\lambda$  corresponding to  $A, A^t$  respectively. We know by exercise 14 of section 5.1 in [FIS03] that  $A$  and  $A^t$  have the same eigenvalues. Thus  $A^t$  is a matrix in which each entry is positive and has eigenvalue  $\lambda$ .

Now notice that because the rows of  $A$  are the columns of  $A^t$  we know  $\nu(A) = \rho(A^t)$  and thus  $|\lambda| = \rho(A^t)$ .

Thus  $A^t$  is a matrix with all positive entries and with an eigenvalue  $|\lambda| = \rho(A^t)$  and by Theorem 5.18 [FIS03], the basis of  $E_{\lambda'} = \{u\}$ , where  $u \in C^n$  is the column vector in which each coordinate contains 1 and  $\lambda = \rho(A^t) = \nu(A)$ .

Thus  $\dim(E_{\lambda'}) = 1$ . But by exercise 13 of 5.2 of [FIS03], we know  $\dim(E_\lambda) = \dim(E_{\lambda'}) = 1$ . Thus, we have shown  $\lambda = \nu(A)$  and  $\dim(E_\lambda) = 1$ .

□

**Corollary 2.2** (Corollary 2 pg. 299). Let  $A \in M_{n \times n}(C)$  be a transition matrix in which each entry is positive and let  $\lambda$  be an eigenvalue of  $A$  such that  $\lambda \neq 1$ . Then  $|\lambda| < 1$  and the eigenspace corresponding to the eigenvalue 1 has dimension 1.

*Proof.* We know by corollary 3 of Theorem 5.16 [FIS03] that if  $\lambda$  is an eigenvalue of a transition matrix, then  $|\lambda| \leq 1$ . Thus if  $|\lambda| \neq 1$ , then  $|\lambda| < 1$ . Suppose  $|\lambda| = 1$ . We know that the sum of each column of  $A$  is 1 and thus  $\nu(A) = 1$  and thus  $|\lambda| = \nu(A)$ . But this means by the previous corollary that  $\lambda = \nu(A) = 1 \neq -1$ . Thus  $\lambda \neq 1$  implies  $|\lambda| < 1$ .

If  $A$  is a transition matrix, then by Theorem 5.17 [FIS03] we know that 1 is an eigenvalue. We also know that if  $A$  is a transition matrix, then given  $u$  as a column vector in which each coordinate equals 1, then  $A^t u = u$ .

Because  $\forall i \ u_i = 1$ , we know

$$1 = u_i = \sum_{j=1}^n A_{ij}^t u_j = \sum_{j=1}^n A_{ij}^t (1) = \sum_{j=1}^n A_{ij}^t = \rho_i(A^t) = \nu_i(A)$$

Then  $\forall i \ \nu_i(A) = 1$  and thus  $\nu(A) = 1$ . Therefore we have an all-positive-entry matrix with  $1 = \lambda = \nu(A)$  and thus by the previous corollary, we know  $\dim(E_\lambda) = 1$ .

□

**Theorem 3** (Theorem 5.19 pg. 298). Let  $A$  be a regular transition matrix and let  $\lambda$  be an eigenvalue of  $A$ . Then

(1)  $|\lambda| \leq 1$ (2) If  $|\lambda| = 1$ , then  $\lambda = 1$  and  $\dim(E_\lambda) = 1$ .

*Proof.* We know (1) was proved in Corollary 3 of Theorem 5.16 [FIS03].

Since  $A$  is regular, we know by definition  $\exists s \in \mathbb{N}_{>0}$  such that  $A^s$  has only positive entries. Moreover  $A^s$  and  $A^{s+1}$  are transition matrices because  $A$  is a transition matrix and the product of transition matrices is a transition matrix. We now split this proof into parts:

- (1) Because  $A$  is a transition matrix and the entries of  $A^s$  are positive, we know the entries of  $A^{s+1} = A^s(A)$  are positive. More specifically,

$$A_{ij}^{s+1} = \sum_{k=1}^n A_{ik}^s A_{kj}$$

and thus because for any given  $i, j$ ,  $\forall 1 \leq k \leq n$   $A_{kj} \geq 0$  and  $A_{ik}^s > 0$  and  $\exists k$  such that  $A_{kj} > 0$  (since the column sums to 1), we can conclude  $A_{ij}^{s+1} > 0$ .

- (2) Suppose  $|\lambda| = 1$ , then we know by problem 15b of section 5.1 in [FIS03] that if  $\lambda$  is a eigenvalue of  $A$ , then  $\lambda^s, \lambda^{s+1}$  are eigenvalues of  $A^s, A^{s+1}$  respectively. Because  $|\lambda| = 1$ , we know  $|\lambda^s| = |\lambda^{s+1}| = |\lambda| = 1$ .
- (3) Because each entry of  $A^s, A^{s+1}$  is positive and both matrices are transition matrices, we know that for any eigenvalues  $\lambda^*$  of  $A^s, A^{s+1}$  such that  $\lambda^* \neq 1$ , then  $|\lambda^*| < 1$ . Thus because  $|\lambda^s| = |\lambda^{s+1}| = 1$ , we can conclude  $\lambda^s = \lambda^{s+1} = 1$  and therefore  $\lambda = 1$ .
- (4) Let  $E_\lambda$  and  $E_\lambda'$  be the eigenspaces of  $A, A^{s+1}$  respectively corresponding to  $\lambda = 1$ . Then  $E_\lambda \subseteq E_\lambda'$  and because  $\dim(E_\lambda') = 1$ , we know  $\dim(E_\lambda) = 1$ .

□

**Corollary 3.1** (Corollary Pg. 300). Let  $A$  be a regular transition matrix that is diagonalizable. Then  $\lim_{m \rightarrow \infty} A^m$  exists.

*Proof.* We know by Theorem 5.14 [FIS03] that if  $A \in M_{n \times n}(C)$  is diagonalizable and has every eigenvalue contained in  $S$ , where  $S = \{\lambda \in C : |\lambda| < 1 \text{ or } |\lambda| = 1\}$ , then  $\lim_{m \rightarrow \infty} A^m$  exists.

Thus because  $A$  is diagonalizable by assumption, we just need to show for all eigenvalues  $\lambda$  of  $A$ ,  $\lambda \in S$ . But we know by Theorem 5.19 that  $\forall$  eigenvalues  $\lambda$ ,  $\lambda = 1$  or  $|\lambda| < 1$  and thus  $\lambda \in S$ . Thus we can conclude that  $\lim_{m \rightarrow \infty} A^m$  exists.

□

The following lemmas use Jordan Canonical form to prove Theorem 5.20.

**Lemma 4** (Modified Exercise 21 of Section 7.2 [FIS03]). Let  $A \in M_{n \times n}(C)$  be a transition matrix. Since  $C$  is an algebraically closed field,  $A$  has a Jordan canonical form  $J$  to which  $A$  is similar. Let  $P$  be an invertible matrix such that  $P^{-1}AP = J$ . Then we have the following:

- (1)  $\|A^m\| \leq 1$  for every positive integer  $m$ .
- (2)  $\exists c > 0$  such that  $\|J^m\| \leq c$  for every positive integer  $m$ .
- (3) Each Jordan block of  $J$  corresponding to the eigenvalue  $\lambda = 1$  is a  $1 \times 1$  matrix.
- (4)  $\lim_{m \rightarrow \infty} A^m$  exists if and only if 1 is the only eigenvalue of  $A$  with absolute value 1.
- (5)

*Proof.* We know that  $A$  has eigenvalue 1 by Theorem 5.17 [FIS03] and the characteristic polynomial of  $A$  splits.

Thus by Theorem 7.4 [FIS03], we know  $\forall 1 \leq i \leq k$ ,  $\dim(K_1)$  is the multiplicity of 1 as an eigenvalue of  $A$ . Additionally, we know  $E_\lambda \subseteq K_\lambda$  by Theorem 7.1.

But if each Jordan block of  $J$  corresponding to the eigenvalue  $\lambda = 1$  is a  $1 \times 1$  matrix, then  $\dim(K_1) = 1$ . But if  $E_\lambda \subsetneq K_1$ , then  $\dim(E_\lambda) = 1$ .

□

**Theorem 5** (Theorem 5.20). Let  $A$  be an  $n \times n$  regular transition matrix. Then:

- (1) The multiplicity of 1 as an eigenvalue of  $A$  is 1.
- (2)  $\lim_{m \rightarrow \infty} A^m$  exists.
- (3)  $L = \lim_{m \rightarrow \infty} A^m$  is a transition matrix.
- (4)  $LA = AL = L$ .

- (5) The columns of  $L$  are identical and equal to the unique probability vector  $v$  that is equal to the eigenvalue 1.  
 (6) For any probability vector  $w$ ,  $\lim_{m \rightarrow \infty} A^m w = v$ .

*Proof.* (1) See exercise 20 of section 7.2

- (2) 1, 5.13, 5.19  
 (3) By theorem 5.15 [FIS03], we know  $L$  is a transition matrix if  $L^t u = u$ , where  $u$  is the column vector where every entry is equal to 1. Due to the equivalence, we can know  $L$  is a transition matrix is  $u^t L = u^t$ . But

$$u^t L = u^t \lim_{m \rightarrow \infty} A^m = \lim_{m \rightarrow \infty} u^t A^m$$

But  $A^m$  is a transition matrix which means  $\forall m, u^t A^m = u^t$  and thus

$$\lim_{m \rightarrow \infty} u^t A^m = \lim_{m \rightarrow \infty} u^t = u^t$$

Thus  $L$  is a transition matrix.

- (4) By Theorem 5.12 [FIS03], we know  $AL = \lim_{m \rightarrow \infty} AA^m = \lim_{m \rightarrow \infty} A^{m+1} = \lim_{m \rightarrow \infty} A^m A = LA$ .  
 But  $\lim_{m \rightarrow \infty} A^{m+1} = \lim_{m \rightarrow \infty} A^m = L$ .  
 Thus  $LA = AL = L$ .

- (5) We know  $AL = L$  by (4). Let  $L_i$  be the  $i$ th column vector of  $L$ . Then because  $AL = L$ , we know  $AL_i = L_i \forall 1 \leq i \leq n$ . Thus  $L_i$  is an eigenvector of  $A$  corresponding to eigenvalue  $\lambda = 1$ .

Additionally, because  $L$  is a transition matrix by (3), we know  $L^t u = u$  and thus  $\forall 1 \leq i \leq n, L_i^t u = u$ . Thus  $\forall 1 \leq i \leq n, L_i$  is a probability vector.

But then using (1), we know that the multiplicity of 1 as an eigenvalue of  $A$  is 1. Thus all the columns of  $L$  have to be a scalar multiple of the same vector. But because every column is a probability vector, they have to be the same in order to satisfy  $\forall 1 \leq i \leq n, L_i^t u = u$ .

need to prove

Thus each column of  $L$  is equal to the unique probability vector  $v$  corresponding to eigenvalue 1.

- (6) Let  $w$  be any probability vector and  $y = \lim_{m \rightarrow \infty} A^m w = Lw$ . We want to show  $y = v$ . If  $y = Lw$ , then by the corollary to Theorem 5.15 [FIS03], we know  $y$  is a probability vector.

Additionally,

$$Ay = A(Lw) = (AL)w = Lw$$

by part (4). But  $Lw = y$  and thus  $Ay = y$ .

Therefore  $y$  is a probability vector and eigenvector of  $A$  corresponding to  $\lambda = 1$ . But  $v$  is the unique probability vector and eigenvector of  $A$  corresponding to  $\lambda = 1$  and thus  $y = v$ .

Thus  $\lim_{m \rightarrow \infty} A^m w = v$ .

□

#### 4. APPLICATIONS

#### REFERENCES

- [FIS03] S.H. Friedberg, A.J. Insel, and L.E. Spence. *Linear Algebra*. Featured Titles for Linear Algebra (Advanced) Series. Pearson Education, 2003.